medRxiv preprint doi: https://doi.org/10.1101/2023.06.13.23291354; this version posted June 19, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Genome-wide meta-analysis of CSF biomarkers in Alzheimer's disease and Parkinson's disease cohorts

Michael Ta[13], MSc, Cornelis Blauwendraat[1], PhD, Tarek Antar[1], BSc, Hampton L Leonard[13], Msc, Andrew B. Singleton[2], PhD, Mike A. Nalls[13], PhD, Hirotaka Iwaki[13*], MD PhD, Alzheimer's Disease Neuroimaging Initiative (ADNI), the Fox Investigation for New Discovery of Biomarkers

1) Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA
2) Center for Alzheimer's and Related Dementias, National Institutes of Health, Bethesda, MD, USA 20892
3) Data Tecnica International, Washington, DC, USA

*Correspondence to: Hirotaka Iwaki (hirotaka.iwaki@mail.nih.gov)
Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health
35 Convent Drive, Bethesda, MD, USA 20892
Phone Number: 301-435-7606
Fax Number: 301-451-7295
Email Address: hirotaka.iwaki@mail.nih.gov

medRxiv preprint doi: https://doi.org/10.1101/2023.06.13.23291354; this version posted June 19, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Abstract

Background: Amyloid beta (Aβ), phosphorylated tau (p-tau), and total tau (t-tau) in cerebrospinal fluid are established biomarkers for Alzheimer's disease (AD). In other neurodegenerative diseases, such as Parkinson's disease (PD), these biomarkers have also been found to be altered, and the molecular mechanisms responsible for these alterations are still under investigation. Moreover, the interplay between these mechanisms and the diverse underlying disease states remains to be elucidated.
Objectives: To investigate genetic contributions to the AD biomarkers and assess the commonality and heterogeneity of the associations per underlying disease status.
Methods: We conducted GWAS for the AD biomarkers on subjects from the Parkinson's Progression Markers Initiative (PPMI), the Fox Investigation for New Discovery of Biomarkers (BioFIND), and the Alzheimer's Disease Neuroimaging Initiative (ADNI) and meta-analyzed with the largest AD GWAS.[7] We tested heterogeneity of associations of interest between different disease statuses (AD, PD, and control).
Results: We observed three GWAS signals: the *APOE* locus for Aβ, the 3q28 locus between *GEMC1* and *OSTN* for p-tau and t-tau, and the 7p22 locus (top hit: rs60871478, an intronic variant for *DNAAF5*, also known as *HEATR2*) for p-tau. The 7p22 locus is novel and co-localized with the brain *DNAAF5* expression. While no heterogeneity from underlying disease status was observed for the above GWAS signals, some disease risk loci suggested disease specific associations with these biomarkers.
Conclusions: Our study identified a novel association at the intronic region of *DNAAF5* associated with increased levels of p-tau across all diseases. We also observed some disease specific genetic associations with these biomarkers.

Introduction

Alzheimer's disease (AD) poses a significant social burden globally.[1, 2] Cerebrospinal fluid (CSF) levels for amyloid beta (Aβ), phosphorylated Tau (p-tau), and total Tau (t-tau) are established AD biomarkers integrated into the NIA-AA (National Institute on Aging – Alzheimer's Association) research framework for Alzheimer's disease.[16] The pathological significance of these biomarkers has been studied, and genome-wide association studies (GWAS) for these biomarkers have identified several GWAS loci associated with AD risk and progression.[6, 7, 9, 19, 26] CSF Aβ levels have been shown to be lower in AD cases whereas levels of CSF p-tau are elevated compared to normal subjects. Interestingly, these biomarkers were also reported to be altered in Parkinson's disease (PD).[13, 17, 37] Previous studies that directly explored the relationship between these CSF biomarkers and PD showed a decrease in levels of both CSF Aβ and p-tau in cases versus control subjects. However, the genetic background of these observations across neurodegenerative diseases have not been well investigated.

In this study, we conducted GWAS on the AD biomarkers of participants from PD focused studies: the Parkinson's Progression Markers Initiative (PPMI)[24] and the Fox Investigation for New Discovery of Biomarkers (BioFIND). We combined these results with the largest GWAS conducted on mixed cohorts for AD and controls.[7] We stratified the analysis on the recruitment study arms of each cohort and assessed the overall genetic contributions across healthy controls and case subjects for AD and PD groups irrespective of clinical phenotype. In addition, we investigated disease specific genetic contributions through the genetic heterogeneity

medRxiv preprint doi: https://doi.org/10.1101/2023.06.13.23291354; this version posted June 19, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

between individuals with PD, AD, and healthy volunteers using the above PD studies as well as the Alzheimer's Disease Neuroimaging Initiative (ADNI). We also assessed the genetic associations with the biomarker changes overtime when longitudinal data were available.

Methods

Participants

PPMI is an ongoing longitudinal observational study with multiple study arms. The current analyses included data from participants with early-stage idiopathic PD but had not yet received medication for PD at enrollment (PPMI_PD), healthy controls (PPMI_HC), those with scans without evidence of dopaminergic deficit but with Parkinsonism (PPMI_SWEDD), and those with prodromal symptoms such as hyposmia, REM sleep behavior disorder, and image confirmed dopaminergic deficit (PPMI_PRODROMAL). This study also included two genetically enriched study arms from PPMI where carriers of any high risk or causal variant for Parkinson's disease (*LRRK2* G2019S, R1441C/G, *GBA1* N409S, L483P, 84GG, and *SNCA* A53T) were recruited. Both carriers with PD less than 7 years from diagnosis (PPMI_GENPD) and unaffected carriers or their 1st degree family members (PPMI_GENUN) were analyzed. BioFIND was a cross-sectional study with two study arms: PD cases in moderately advanced stages (BioFIND_PD) and healthy controls (BioFIND_HC). Both of these study arms were included in this study. The protocols for these studies can be obtained from the Michael J. Fox Foundation for Parkinson's research (https://www.michaeljfox.org). We also included participants from the Alzheimer's Disease Neuroimaging Initiative (ADNI, https://adni.loni.usc.edu). Based on their last diagnosis, these participants were stratified as either having dementia, (ADNI-Dementia), mild cognitive impairment (ADNI-MCI), or normal cognition (ADNI-NC). While clinical diagnosis of AD can change over time, 97% of the ADNI-Dementia were "probable" AD by NINCDS/ADRDA criteria according to the last record when available. Clinical data of the study participants such as disease status, age, sex, age at diagnosis were obtained from the study websites on December 12th, 2021. For a summary of the study design and data used across disease statuses, please refer to **Supplemental Figure 1**. Descriptive statistics for each study are available in **Supplementary Table 1**.

CSF biomarkers

For ADNI and PPMI samples, CSF concentrations of Amyloid-β 1to42 (Aβ), total tau (t-tau), and phosphorylated tau at the threonine 181 position (p-tau) were measured using Elecsys electro-chemiluminescence immunoassays on the cobas e 601 analysis platform (Roche Diagnostics).[28] For BioFIND samples, these biomarkers were measured by INNO-BIA AlzBio3 immunoassay.[23] The detailed procedures and quality control process are summarized on the study websites. For GWAS analyses, biomarker values were log transformed and centered at zero to be compatible with existing summary statistics.

Genetic data

We used the whole genome sequencing data (WGS) provided by the ADNI repository and the AMP-PD project.[15] The samples were sequenced (30x or more coverage) and underwent the GATK best practices workflow. Additional details regarding quality control are provided on the study websites. In this analysis we used PASS filtered variants and analyzed only the

medRxiv preprint doi: https://doi.org/10.1101/2023.06.13.23291354; this version posted June 19, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

participants with European ancestry because of insufficient power to analyze non-European ancestry groups. The ancestry was confirmed by being within +/- 6SD of the first two principal components of the European samples (CEU and TSI) in HapMap3 panel.[12] We also excluded related individuals closer than 2nd degree relatives from the analysis.

Summary statistics from the prior GWAS

We requested the summary statistics from the largest GWAS for Aβ, t-Tau and p-Tau from NIAGADS (https://www.niagads.org/) data repository (NG00055).[7] The GWAS included 3,146 individuals with and without dementia from nine different studies conducted at the Charles F. and Joanne Knight Alzheimer's Disease Research Center (Knight ADRC), Saarland University in Homburg/Saar, Germany (HB), Mayo Clinic (MAYO), Skåne University Hospital, Sweden (SWEDEN), Perelman School of Medicine at the University of Pennsylvania (UPENN), and the University of Washington (UW) as well as Alzheimer's Disease Neuroimaging Initiative (ADNI1 and ADNI2) and Predictors of Cognitive Decline Among Normal Individuals (BIOCARD). The CSF biomarkers were log-transformed and centered per study followed by a single-stage association test adjusted for age, sex, measurement platform, and the first two principal components.

Analysis

Cohort-strata-level GWAS for the AD biomarkers were conducted. For the BioFIND study, we fit a linear regression model for additive allele effect using the cross-sectional data. For the other studies where longitudinal data were available, we used the GALLOP algorithm to approximate the linear mixed effects model for both the additive allele effect (cross-sectional associations) and the additive *allele x time* interaction (longitudinal associations). This algorithm provides equivalent solutions to a linear mixed effects model in a computationally efficient way.[29] In both models, we adjusted for age, sex, and the first two principal components (PC1-PC2). For the GALLOP model, we further adjusted for time from the baseline measurement, interactions between time and PCs (PC1-PC2), and a random intercept and random slope for each individual. Our primary analysis was to meta-analyze the cross-sectional results with the previously reported summary statistics from the largest CSF AD biomarker GWAS[7] to identify the across-disease genetic contributions for these biomarkers. We also meta-analyzed the longitudinal associations to see if there are any genome-wide significant loci associated with the biomarker change over time.

In the meta-analysis, all of the variants with a minor allele count less than 5 or minor allele frequency <1% among the individual studies, not reported in more than 2 study arms, or failed for heterogeneity assessment (p-value for the test of heterogeneity < 0.05 or I2 > 80%) were removed for from the "overall" genetic assessment for the biomarkers. For novel genome-wide significant loci, we further assessed co-localization with brain eQTL (expression quantitative trait loci)[3] and blood eQTL[33] using LocusCompare.[21]

For those SNPs reported in the previous study by Deming et al, we assessed if there was any evidence of heterogeneity between different disease states. We meta-analyzed GWAS from PPMI-PD, PPMI-GENPD, BioFIND-PD to compose "PD" GWAS summary statistics. Similarly, we meta-analyzed ADNI-CN, BioFIND-HC, and PPMI-HC and PPMI-GENUN to generate "HC" GWAS results. For AD, we used the ADNI-Dementia GWAS summary statistics. The heterogeneity between these disease specific GWAS results were assessed using their I^2 statistics.

medRxiv preprint doi: https://doi.org/10.1101/2023.06.13.23291354; this version posted June 19, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Finally, we assessed if disease specific or nonspecific genetic associations with these CSF biomarkers exist for known risk loci associated with either PD or AD.[4, 22] For this targeted analysis, the significance level was set at the false positive rate (q-value) of 0.05 adjusting for the number of loci to be tested.

All the statistical analyses and drawings were executed using Plink version 2.0 alpha,[25] R version 3.6 and Python version 3.8. Meta-analyses were conducted using METAL software with an inverse variance weighted method with the genomic control correction applied.[36] The analysis scripts are available at https://github.com/NIH-CARD/biomarker_longGWAS. The data were obtained from ADNI and AMP-PD.

Results

We used data from 61 GWAS in the main meta-analyses, including previously published results: 34 GWAS for cross-sectional components and 27 GWAS for longitudinal components. The genomic inflation factors of the GWAS were reasonable, with most around 1.0 except for three between 1.1-1.3 (Supplemental Table 2). The inflation was accounted for in the meta-analysis phase by the genomic control function in METAL.

In the meta-analysis, we observed three GWAS signals in the cross-sectional component: the *APOE* locus for Aβ, the 3q28 locus between *GEMC1* and *OSTN* for p-tau and t-tau, and the 7p22 locus (top hit: rs60871478, an intron variant for *DNAAF5*, also known as *HEATR2*) for p-tau (Table 1, Figure 1, Supplemental Figure 2 and 3). No genome-wide signals were identified in the longitudinal components (Supplemental Figure 5). The rs769449 variant on the APOE locus was identified to be significant for all three CSF biomarkers. For rs35055419, a significant association was seen for p-tau and t-tau. Both of these variants were previously reported by Cruchaga et. al.

The chromosome 7 locus in the *DNAAF5* gene region was not previously reported and showed genome-wide significant P-value for p-tau P=1.97E-8 and a sub genome-wide significant P-value for t-tau P=5.82E-7. When exploring the potential causal gene in this region we identified eQTLs for *DNAAF5* that were well-colocalized with the GWAS signals in both the brain and blood (Supplemental Figure 4). No other genes in this locus showed any colocalization between the QTL and the GWAS signal.

Disease specific genetic associations with CSF biomarkers was assessed on the significant loci reported by Deming et al. between AD and PD related dementia. A stratified meta-analysis by disease state suggested heterogeneity was present for one of the non-replicated associations between rs12961169 (*CTDP1*) and p-tau. The ADNI-Dementia group showed a negative association (p=0.00182) that was not observed in the PD and the HC groups (Supplemental Figure 6 and 7). No evidence of heterogeneity for the other non-replicated associations was seen. One reported locus at 1p32.3 for Aβ (rs185031519, very rare) was not identified in the current analysis.

Likewise, for known AD and PD risk associated loci on CSF biomarkers, the disease stratified meta-analysis was used to identify the genetic similarities and differences across the dementias and control groups. The *APOE* e4 tagging allele was associated with lower CSF Aβ regardless

medRxiv preprint doi: https://doi.org/10.1101/2023.06.13.23291354; this version posted June 19, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

of the disease status (Table 2). AD risk increasing allele, rs6586028_T (*TSPAN14*) was associated with the increasingly lower CSF Ab over time in the PD group but not in the ADNI-Dementia and the HC groups, although there was not enough evidence suggesting heterogeneity among these results. The PD risk increasing allele, rs7134559_C (12q13.11*)* was associated with lower p-tau in a disease status non-specific way.

Discussion

In this study, we conducted GWAS on CSF levels of three known AD biomarkers. We used data from PD and AD studies and assessed multiple diseases and stages of progression. We replicated two genetic loci from the previous largest AD biomarker study[7] (*APOE* and *GEMC1*) and identified a new locus at 7p22 that reached genome-wide significance in association with p-tau. *APOE* was previously shown to have significant associations with CSF biomarkers and is a genetic risk factor for late-onset sporadic AD[5]. In PD, carriers of the *APOE*-ε4 allele were found to have both quicker cognitive decline compared to non-carriers and an increased risk of progression to dementia.[30] Prior studies have found rs9877502 on the 3q28 locus between *GEMC1* and *OSTN* to be associated with higher CSF tau levels and identified a risk variant (rs1316356) for AD that is in linkage disequilibrium with this SNP.[6, 7] Additionally, *GEMC1* has been recently reported to be a key molecule in multiciliated cell differentiation.[20] In the brain, these cells are involved in maintaining homeostasis and neurogenesis. The new 7p22 locus, rs60871478, was associated with increased levels of CSF p-tau regardless of the disease status and colocalized well with *DNAAF5* (also known as *HEATR2*). When exploring the potential causal gene in this region we identified that there was good colocalization between the GWAS signal and blood and brain eQTL data for *DNAAF5* and no correlation was seen with other genes in the locus.

This gene encodes the protein Dynein Axonemal Assembly Factor 5 and is essential for the pre-assembly or stability of axonemal dynein. A missense mutation in *DNAAF5* was identified in a whole-exome sequencing study of a family with primary ciliary dyskinesia, a rare autosomal recessive disease that presents with neonatal respiratory distress, sinopulmonary disease, otitis media, male infertility, and left-right laterality defects. The affected individuals showed a malfunction in airway epithelial cells.[11] The gene is expressed ubiquitously across all tissues however the link between p-tau and the gene is unclear.

To assess the clinical consequence of the variant, we conducted ad-hoc analyses testing associations of this locus with age at onset and MMSE in the ADNI cohort (Supplemental Table 3). The results suggest the loci is associated with MMSE scores in the ADNI-Dementia group (-1.07±0.48, p-val: 0.025). A similar association was shown in the ADNI-CN group with a smaller magnitude of effect (-0.34±0.17, p-val: 0.021) . This variant may play a role in cognitive decline related to increased tau pathology with underlying AD pathology, but the sample size is not large enough to adjust for multiple testing and further evaluation with a larger cohort would be required. In addition, biological evaluation such as cellular or animal Alzheimer models may provide more information regarding progression and dementia risk associated with the locus.

Six previously reported cross-sectional associations with CSF biomarkers in AD were not replicated in this study. Of these, rs12961169 (*CTDP1*) showed high heterogeneity across diseases. For the others, it may in part be due to differences in the study designs. The previous study was focused on identifying AD related loci using the biomarkers as endophenotypes. So, they conducted one-stage GWAS without adjusting for the disease status.

medRxiv preprint doi: https://doi.org/10.1101/2023.06.13.23291354; this version posted June 19, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Interestingly, the PD risk increasing loci rs7134559 (*SCAF11*) was associated with the lower CSF p-tau. A recent study reported that these biomarkers were indeed lower in PD when compared with HC in contrast to the generally higher CSF p-tau in AD.[13] The reason for the difference in these biomarker profiles between the two diseases is unknown, but there are many reports suggesting the influence of AD pathology to the PD pathology or vice versa.[14, 27, 35] The current observation may be associated with some interaction between the two disease mechanisms.

Longitudinal changes are smaller compared to baseline differences in biomarker levels between disease states. Reported heterogeneity on CSF biomarker trajectories has been observed in AD risk allele carriers prior to and after the onset of dementia symptoms.[8] The need for a stratified analysis by dementia progression in addition to larger cohort sizes might improve detection of longitudinal genetic contributions.

By integrating data from PD studies, we were able to expand the knowledge of genetic-biomarker relationships that were mainly derived from AD studies previously. For some SNPs, we could differentiate disease specific and non-specific genetic associations. Admittingly, the sizes of this study, especially in regards to the disease specific GWAS, were still small. Additional data is needed for more effective analyses in particular large datasets from diverse ancestries with longitudinal measures available. Nevertheless, we believe that the current approach would be useful to investigate underlying AD mechanisms modified by different disease status. Another limitation of this study is the potential misdiagnosis of AD and PD because the clinical diagnosis is not always accurate. In particular, misdiagnosis would affect the heterogeneity assessment. Access to additional biomarkers and increasing study sizes are both important to overcome this problem.

We also like to note that this study may have not fully accounted for the complex relationships between Aβ, p-tau, and t-tau. These biomarkers are recognized as endophenotypes, and they are highly sensitive and specific in differentiating AD and controls.[5, 10] However, the clinical significance of these biomarkers is not equal. First, they are thought to represent different pathological processes related to AD: low CSF Aβ for aggregation of Aβ, high CSF pTau for neurofibrillary tangle (NFT) formation, and high CSF tTau for neurodegeneration.[32] The timing of deviation from normal is also different, as the decrease of Aβ is observed earlier than the increase of CSF p-Tau and t-Tau.[10] Additionally, the increase of tTau and pTau in the early stage of AD may be associated with faster progression of disease.[31, 34] Moreover, there is a study that reported the level of p-Tau, supposedly reflecting tangle pathology, was more closely associated with amyloid PET than with tau PET.[18] To further investigate these complex relationships between the biomarkers, GWAS on various stratifications, such as disease status, other biomarker status and imaging status, should provide important information to untangle these relationships.

In conclusion, we analyzed the CSF AD biomarker from the AD and the PD studies. We identified three associations across disease status including one novel genome-wide significant locus and also observed some associations suggesting the disease specific modifications of these biomarkers at known risk loci.

Acknowledgement

medRxiv preprint doi: https://doi.org/10.1101/2023.06.13.23291354; this version posted June 19, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

The research team would like to thank the participants and individuals involved in the organization and collection of study materials.

This work was supported by the Center for Alzheimer's and Related Dementias, within the Intramural Research Program of the National Institute on Aging and the National Institute of Neurological Disorders and Stroke, National Institutes of Health, Department of Health and Human Services. Project numbers Z01-AG000957-19 and Z01-AG000534-03. This work utilized the computational resources of the NIH HPC Biowulf cluster. (http://hpc.nih.gov).

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.;Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.;Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Data used in the preparation of this article were obtained from the Parkinson's Progression Markers Initiative (PPMI) database (www.ppmi-info.org/access-data-specimens/download-data). For up-to-date information on the study, visit www.ppmi-info.org. PPMI – a public-private partnership – is funded by the Michael J. Fox Foundation for Parkinson's Research and funding partners, including: 4D Pharma; AbbVie Inc.; AcureX Therapeutics; Allergan; Amathus Therapeutics; Aligning Science Across Parkinson's (ASAP); Avid Radiopharmaceuticals; Bial Biotech; Biogen; BioLegend; Bristol Myers Squibb; Calico Life Sciences LLC; Celgene Corporation; DaCapo Brainscience; Denali Therapeutics; The Edmond J. Safra Foundation; Eli Lilly and Company; GE Healthcare; GlaxoSmithKline; Golub Capital; Handl Therapeutics; Insitro; Janssen Pharmaceuticals; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC; Neurocrine Biosciences; Pfizer Inc.; Piramal Imaging; Prevail Therapeutics; F. Hoffmann La Roche Ltd and its affiliated company Genentech Inc.; Sanofi Genzyme; Servier; Takeda Pharmaceutical Company; Teva Neuroscience, Inc.; UCB; Vanqua Bio; Verily Life Sciences; Voyager Therapeutics, Inc.; and Yumanity Therapeutics, Inc.

BioFIND is funded by The Michael J. Fox Foundation for Parkinson's Research and the National Institute Neurological Disorders and Stroke.

Conflict of Interest

M.T., H.L.L., M.A.N. and H.I.'s participation in this project was part of a competitive contract awarded to Data Tecnica International LLC by the National Institutes of Health to support open

medRxiv preprint doi: https://doi.org/10.1101/2023.06.13.23291354; this version posted June 19, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

science research. M.A.N. also currently serves on the scientific advisory board for Character Bio Inc. and is an advisor to Neuron23 Inc.

medRxiv preprint doi: https://doi.org/10.1101/2023.06.13.23291354; this version posted June 19, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

References

1. † (2018) Global, regional, and national burden of Parkinson's disease, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol 17(11):939–953

2. † (2019) Global, regional, and national burden of Alzheimer's disease and other dementias, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol 18(1):88–106

3. Allen M, Carrasquillo MM, Funk C, et al (2016) Human whole genome genotype and transcriptome data for Alzheimer's and other neurodegenerative diseases. Sci Data 3:160089

4. Bellenguez C, Küçükali F, Jansen IE, et al (2022) New insights into the genetic etiology of Alzheimer's disease and related dementias. Nat Genet 54(4):412–436

5. Blennow K (2017) A Review of Fluid Biomarkers for Alzheimer's Disease: Moving from CSF to Blood. Neurol Ther 6(Suppl 1):15–24

6. Cruchaga C, Kauwe JSK, Harari O, et al (2013) GWAS of cerebrospinal fluid tau levels identifies risk variants for Alzheimer's disease. Neuron 78(2):256–268

7. Deming Y, Li Z, Kapoor M, et al (2017) Genome-wide association study identifies four novel loci associated with Alzheimer's endophenotypes and disease modifiers. Acta Neuropathol 133(5):839–856

8. Fagan AM, Xiong C, Jasielec MS, et al (2014) Longitudinal change in CSF biomarkers in autosomal-dominant Alzheimer's disease. Sci Transl Med 6(226):226ra30

9. Han M-R, Schellenberg GD, Wang L-S, Alzheimer's Disease Neuroimaging Initiative (2010) Genome-wide association reveals genetic effects on human Aβ42 and τ protein levels in cerebrospinal fluids: a case control study. BMC Neurol 10:90

10. Hansson O, Seibyl J, Stomrud E, et al (2018) CSF biomarkers of Alzheimer's disease concord with amyloid-β PET and predict clinical progression: A study of fully automated immunoassays in BioFINDER and ADNI cohorts. Alzheimers Dement 14(11):1470–1481

11. Horani A, Druley TE, Zariwala MA, et al (2012) Whole-exome capture and sequencing identifies HEATR2 mutation as a cause of primary ciliary dyskinesia. Am J Hum Genet 91(4):685–693

12. International HapMap 3 Consortium, Altshuler DM, Gibbs RA, et al (2010) Integrating common and rare genetic variation in diverse human populations. Nature 467(7311):52–58

13. Irwin DJ, Fedler J, Coffey CS, et al (2020) Evolution of Alzheimer's Disease Cerebrospinal Fluid Biomarkers in Early Parkinson's Disease. Ann Neurol 88(3):574–587

14. Irwin DJ, Lee VM-Y, Trojanowski JQ (2013) Parkinson's disease dementia: convergence of α-synuclein, tau and amyloid-β pathologies. Nat Rev Neurosci 14(9):626–636

15. Iwaki H, Leonard HL, Makarious MB, et al (2020) Accelerating Medicines Partnership: Parkinson's Disease. Genetic Resource. medRxiv 2020.11.19.20235192

medRxiv preprint doi: https://doi.org/10.1101/2023.06.13.23291354; this version posted June 19, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

16. Jack CR Jr, Bennett DA, Blennow K, et al (2018) NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. Alzheimers Dement 14(4):535–562

17. Kang J-H, Irwin DJ, Chen-Plotkin AS, et al (2013) Association of cerebrospinal fluid β-amyloid 1-42, T-tau, P-tau181, and α-synuclein levels with clinical features of drug-naive patients with early Parkinson disease. JAMA Neurol 70(10):1277–1287

18. Kester MI, van der Vlies AE, Blankenstein MA, Pijnenburg YAL, van Elk EJ, Scheltens P, van der Flier WM (2009) CSF biomarkers predict rate of cognitive decline in Alzheimer disease. Neurology 73(17):1353–1358

19. Kim S, Swaminathan S, Shen L, et al (2011) Genome-wide association study of CSF biomarkers Abeta1-42, t-tau, and p-tau181p in the ADNI cohort. Neurology 76(1):69–79

20. Lalioti M-E, Arbi M, Loukas I, et al (2019) GemC1 governs multiciliogenesis through direct interaction with and transcriptional regulation of p73. J Cell Sci. doi: 10.1242/jcs.228684

21. Liu B, Gloudemans MJ, Rao AS, Ingelsson E, Montgomery SB (2019) Abundant associations with gene expression complicate GWAS follow-up. Nat Genet 51(5):768–769

22. Nalls MA, Blauwendraat C, Vallerga CL, et al (2019) Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. Lancet Neurol 18(12):1091–1102

23. Olsson A, Vanderstichele H, Andreasen N, De Meyer G, Wallin A, Holmberg B, Rosengren L, Vanmechelen E, Blennow K (2005) Simultaneous measurement of beta-amyloid(1-42), total tau, and phosphorylated tau (Thr181) in cerebrospinal fluid by the xMAP technology. Clin Chem 51(2):336–345

24. Parkinson Progression Marker Initiative (2011) The Parkinson Progression Marker Initiative (PPMI). Prog Neurobiol 95(4):629–635

25. Purcell S, Neale B, Todd-Brown K, et al (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81(3):559–575

26. Ramirez A, van der Flier WM, Herold C, et al (2014) SUCLG2 identified as both a determinator of CSF Aβ1-42 levels and an attenuator of cognitive decline in Alzheimer's disease. Hum Mol Genet 23(24):6644–6658

27. Sengupta U, Kayed R (2022) Amyloid β, Tau, and α-Synuclein aggregates in the pathogenesis, prognosis, and therapeutics for neurodegenerative diseases. Prog Neurobiol 214:102270

28. Shaw LM, Waligorska T, Fields L, et al (2018) Derivation of cutoffs for the Elecsys amyloid β (1-42) assay in Alzheimer's disease. Alzheimers Dement 10:698–705

29. Sikorska K, Lesaffre E, Groenen PJF, Rivadeneira F, Eilers PHC (2018) Genome-wide Analysis of Large-scale Longitudinal Outcomes using Penalization -GALLOP algorithm. Sci Rep 8(1):6815

30. Szwedo AA, Dalen I, Pedersen KF, et al (2022) GBA and APOE Impact Cognitive Decline in Parkinson's Disease: A 10-Year Population-Based Study. Mov Disord 37(5):1016–1027

medRxiv preprint doi: https://doi.org/10.1101/2023.06.13.23291354; this version posted June 19, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

31. Therriault J, Vermeiren M, Servaes S, et al (2023) Association of Phosphorylated Tau Biomarkers With Amyloid Positron Emission Tomography vs Tau Positron Emission Tomography. JAMA Neurol 80(2):188–199

32. Villemagne VL, Pike KE, Chételat G, et al (2011) Longitudinal assessment of Aβ and cognition in aging and Alzheimer disease. Ann Neurol 69(1):181–192

33. Võsa U, Claringbould A, Westra H-J, et al (2021) Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. Nat Genet 53(9):1300–1310

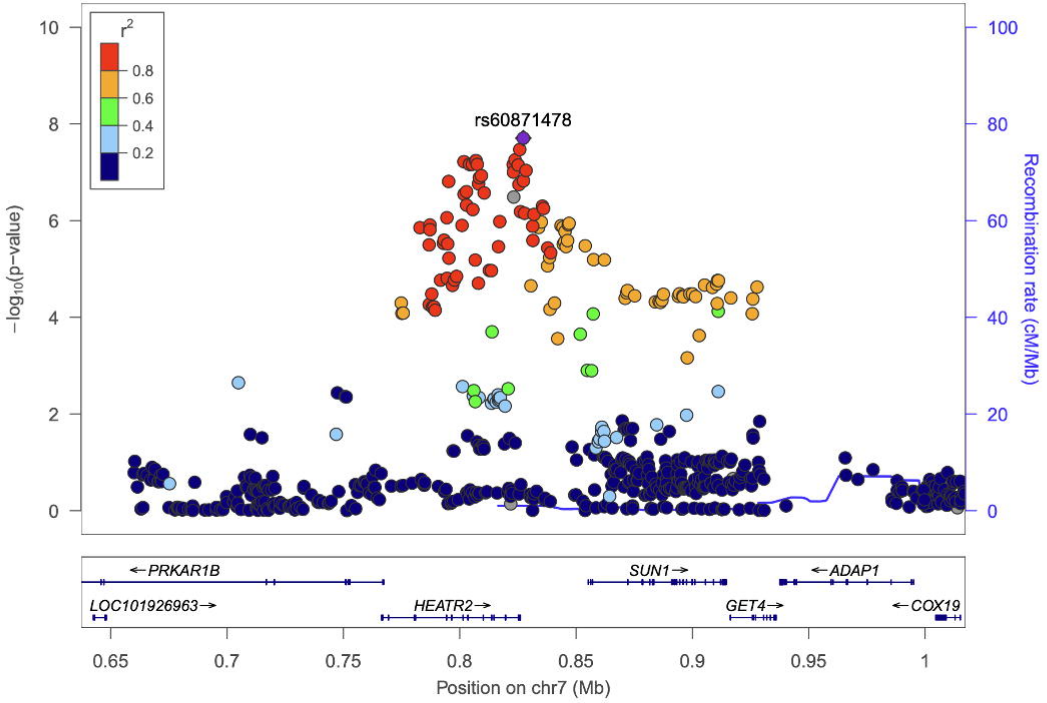
34. Wattmo C, Blennow K, Hansson O (2020) Cerebro-spinal fluid biomarker levels: phosphorylated tau (T) and total tau (N) as markers for rate of progression in Alzheimer's disease. BMC Neurol 20(1):10

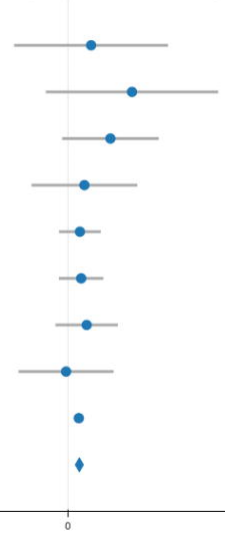
35. Wennberg AM, Whitwell JL, Tosakulwong N, et al (2019) The influence of tau, amyloid, alpha-synuclein, TDP-43, and vascular pathology in clinically normal elderly individuals. Neurobiol Aging 77:26–36

36. Willer CJ, Li Y, Abecasis GR (2010) METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 26(17):2190–2191

37. Zhang J, Mattison HA, Liu C, et al (2013) Longitudinal assessment of tau and amyloid beta in cerebrospinal fluid of Parkinson disease. Acta Neuropathol 126(5):671–682

Figure 1. rs60871478 locus for log CSF p-tau and forest plots



log CSF p-Tau Cross-Sectional - rs60871478 (chr7:787688:A:G)

Study	N (obs)	Mean [95% C.I.]	P-value
BioFIND-PD	84.0 (84.0)	0.07 [-0.17; 0.32]	0.557
BioFIND-Healthy-Control	62.0 (62.0)	0.20 [-0.07; 0.48]	0.151
PPMI-SWEDD	56.0 (174.0)	0.14 [-0.02; 0.29]	0.085
PPMI-Prodromal	53.0 (230.0)	0.05 [-0.12; 0.22]	0.542
PPMI-PD	367.0 (1878.0)	0.04 [-0.03; 0.10]	0.259
PPMI-Genetic-Cohort-Unaffected	280.0 (666.0)	0.04 [-0.03; 0.11]	0.243
PPMI-Healthy-Control	172.0 (893.0)	0.06 [-0.04; 0.16]	0.240
PPMI-Genetic-Cohort-PD	169.0 (434.0)	-0.01 [-0.16; 0.15]	0.939
Deming et. al.	2810.0 (2810.0)	0.03 [0.02; 0.05]	4.505E-07
METAL meta-analysis HetISq=0.00; HetPVal=0.872	4053.0 (4053.0)	0.04 [0.02; 0.05]	1.97e-8